

## PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

## NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Assistant Commissioner for Patents  
United States Patent and Trademark  
Office  
Box PCT  
Washington, D.C. 20231  
ÉTATS-UNIS D'AMÉRIQUE

in its capacity as elected Office

Date of mailing:

04 November 1999 (04.11.99)

International application No.:

PCT/GB99/01278

Applicant's or agent's file reference:

A25587 WO

International filing date:

23 April 1999 (23.04.99)

Priority date:

27 April 1998 (27.04.98)

Applicant:

CORLEY, Stephen, Leslie et al

1. The designated Office is hereby notified of its election made:



in the demand filed with the International preliminary Examining Authority on:

20 September 1999 (20.09.99)



in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was

was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO  
34, chemin des Colombettes  
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

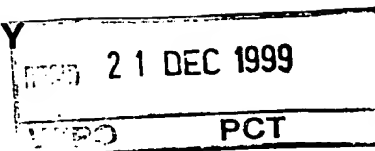
Authorized officer:

J. Zahra

Telephone No.: (41-22) 338.83.38

**PATENT COOPERATION TREATY**

**PCT**





**INTERNATIONAL PRELIMINARY EXAMINATION REPORT**

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference <b>A25587 WO</b>		<b>FOR FURTHER ACTION</b> See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. <b>PCT/GB99/01278</b>	International filing date (day/month/year) <b>23/04/1999</b>	Priority date (day/month/year) <b>27/04/1998</b>
International Patent Classification (IPC) or national classification and IPC <b>G06F17/30</b>		
Applicant <b>BRITISH TELECOMMUNICATIONS PUBLIC LIMITED... et al</b>		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 4 sheets, including this cover sheet.  
  
☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).  
  
 These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:
  - I ☒ Basis of the report
  - II ☐ Priority
  - III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
  - IV ☐ Lack of unity of invention
  - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
  - VI ☐ Certain documents cited
  - VII ☒ Certain defects in the international application
  - VIII ☐ Certain observations on the international application

Date of submission of the demand  <b>20/09/1999</b>	Date of completion of this report  <b>17.12.1999</b>
Name and mailing address of the international preliminary examining authority:   <b>European Patent Office</b> <b>D-80298 Munich</b> <b>Tel. +49 89 2399 - 0 Tx: 523656 epmu d</b> <b>Fax: +49 89 2399 - 4465</b>	Authorized officer  <b>Oestergaard, M</b>  Telephone No. +49 89 2399 2551  

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT**

International application No. PCT/GB99/01278

**I. Basis of the report**

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

**Description, pages:**

1-17 as originally filed

**Claims, No.:**

1-5 as originally filed

**Drawings, sheets:**

1-5 as originally filed

2. The amendments have resulted in the cancellation of:

- ☐ the description, pages:  
☐ the claims, Nos.:  
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT**

International application No. PCT/GB99/01278

**V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

**1. Statement**

Novelty (N)	Yes:	Claims	1-5
	No:	Claims	
Inventive step (IS)	Yes:	Claims	1-5
	No:	Claims	
Industrial applicability (IA)	Yes:	Claims	1-5
	No:	Claims	

**2. Citations and explanations**

**see separate sheet**

**VII. Certain defects in the international application**

The following defects in the form or contents of the international application have been noted:

**see separate sheet**

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT - SEPARATE SHEET**

---

International application No. PCT/GB99/01278

Section V:

- 1 The nearest prior art is found to be D1 = IBM Technical Disclosure Bulletin, vol. 36, No. 1, January 1993, pages 82-83. This document discloses a library list to allow a user to maintain and use SQL statements that were previously created and saved. Otherwise, when a user creates a SQL statement, it is either lost after execution or stored separately as an individual object.

In addition to storing a query as in document D1, in the method as defined by claim 1 storing of the results of the query can be done. If it is decided in a selection that there may be relevant query results for a newly submitted query, then an actual query search does not have to be performed, thus enhancing performance of the database system compared to prior art disclosed by D1.

Compared to D1 the current set of claims appears to be novel. When comparing with the available prior art the current set of claims also appears to include an inventive step. The claims appear to be industrially applicable.

Section VII:

- 2 The nearest prior art D1 should preferably have been reflected in the description and the preamble of the independent claim as required by Rule 5.1(a)(ii and iii) PCT.

## PCT

## INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference <b>A25587 WO</b>	<b>FOR FURTHER ACTION</b> see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. <b>PCT/GB 99/ 01278</b>	International filing date (day/month/year) <b>23/04/1999</b>	(Earliest) Priority Date (day/month/year) <b>27/04/1998</b>
Applicant <b>BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY</b>		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 3 sheets.



It is also accompanied by a copy of each prior art document cited in this report.

## 1. Basis of the report

- a. With regard to the **language**, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.



the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

- b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing:



contained in the international application in written form.



filed together with the international application in computer readable form.



furnished subsequently to this Authority in written form.



furnished subsequently to this Authority in computer readable form.



the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.



the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. ☐ **Certain claims were found unsearchable** (See Box I).

3. ☐ **Unity of invention is lacking** (see Box II).

4. With regard to the **title**,

the text is approved as submitted by the applicant.



the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

the text is approved as submitted by the applicant.



the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No.

3



as suggested by the applicant.



None of the figures.



because the applicant failed to suggest a figure.



because this figure better characterizes the invention.

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 99/01278

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	"SQL STATEMENT LIBRARY LIST" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 36, no. 1, 1 January 1993 (1993-01-01), pages 82-83, XP000333783 ISSN: 0018-8689 the whole document ---	1
A	"STORING QUERY OBJECTS BY TYPE IN A RELATIONAL DATABASE" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 32, no. 4B, 1 September 1989 (1989-09-01), page 14 XP000067092 ISSN: 0018-8689 the whole document --- -/--	1

☒ Further documents are listed in the continuation of box C.☐ Patent family members are listed in annex.

## \* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&amp;" document member of the same patent family

Date of the actual completion of the international search

2 August 1999

Date of mailing of the international search report

09/08/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Deane, E

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 99/01278

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	"STORE SEARCH RESULTS FOR LATER REFERENCE" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 32, no. 108, 1 March 1990 (1990-03-01), page 210 XP000097865 ISSN: 0018-8689 the whole document ----	5
A	PATENT ABSTRACTS OF JAPAN vol. 098, no. 005, 30 April 1998 (1998-04-30) & JP 10 011339 A (HITACHI LTD), 16 January 1998 (1998-01-16) abstract & US 5 905 983 A abstract -----	5



### Information on patent family members

PCT/GB 99/01278

18-05-1999



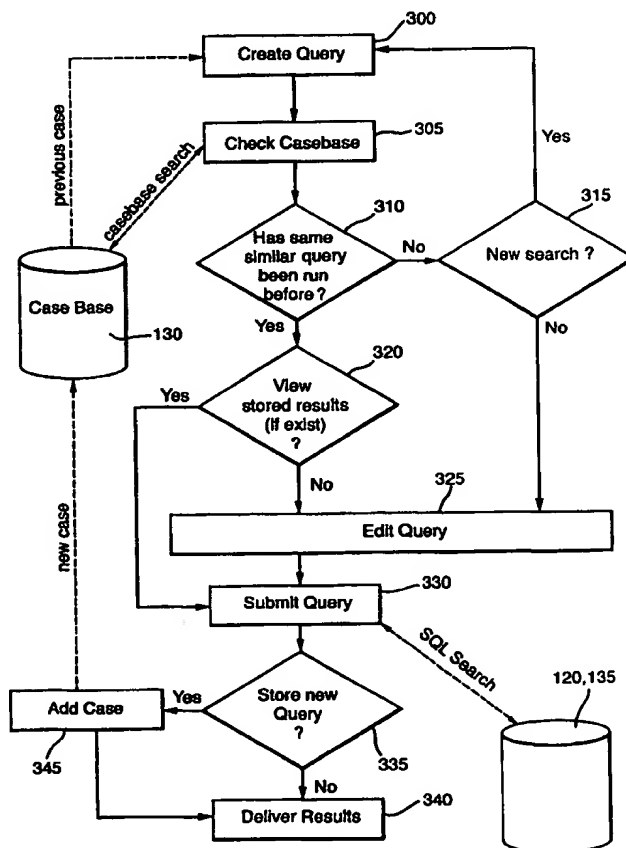
## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>G06F 17/30</b>		A1	(11) International Publication Number: <b>WO 99/56227</b>
			(43) International Publication Date: 4 November 1999 (04.11.99)
(21) International Application Number: PCT/GB99/01278		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 23 April 1999 (23.04.99)			
(30) Priority Data: 9808988.1 27 April 1998 (27.04.98) GB 98307123.4 4 September 1998 (04.09.98) EP			
(71) Applicant (for all designated States except US): BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY [GB/GB]; 81 Newgate Street, London EC1A 7AJ (GB).			
(72) Inventors; and (75) Inventors/Applicants (for US only): CORLEY, Stephen, Leslie [GB/GB]; 7 Browns Grove, Kesgrave, Ipswich, Suffolk IP5 2GP (GB). FAGAN, Michael [GB/GB]; 2 Lime Tree Farm, Lower-Hacheston, Woodbridge, Suffolk IP13 0PB (GB).			
(74) Agent: DUTTON, Erica, Lindley, Graham; BT Group Legal Services, Intellectual Property Dept., Holborn Centre, 8th floor, 120 Holborn, London EC1N 2TE (GB).		Published With international search report.	

(54) Title: DATABASE ACCESS TOOL

## (57) Abstract

Embodiments of the invention alleviate problems relating to databases access, particularly when the database is very large and complex such as in the marketing domain. This domain is characterised by very large databases containing many Gigabytes of customer related information. Many different types of users can need to query a database and in large businesses. Different users may have different abilities or familiarity with database access. Embodiments of the present invention provide a database access tool which makes database access easier for users having different abilities by storing previous queries and reusing them in response to subsequent queries.



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

DATABASE ACCESS TOOL

The present invention relates to querying databases.

Databases can be very large and complex. An example is the marketing  
5 domain. This domain is characterised by very large databases containing many  
Gigabytes of customer related information. Many different types of users can need to  
query a database. In large businesses, there are usually many roles associated with  
accessing data. These roles include people who specify queries, those that test  
them, tuners who carry out optimisation and then those who execute the queries.

10 Queries to such databases need to be carefully structured and query  
languages have developed, such as Structured Query Language (SQL). The SQL  
language consists of a set of facilities for defining, manipulating and controlling data in  
a relational database. SQL, pronounced Sequel, was originally an abbreviation for  
Structured Query Language. SQL is essentially two languages, a data definition  
15 language (DDL), for declaring structures and integrity constraints of SQL databases,  
and a data manipulation language (DML), for declaring the database procedures and  
executable statements of a specific database application program. It is both a formal  
standard and an industry standard. SQL was formally "standardised" by the American  
National Standards Institute (ANSI) in 1986 as ANS X3.135.1986 (normally known as  
20 SQL86.) The ANSI SQL standard was fully adopted by ISO in 1987 as IS 9075:1987.

SQL remains complex and the creation of a database query for a database of  
any complexity is still a specialist and often time consuming activity.

According to embodiments of the present invention, there is provided:

- (i) means for constructing database queries;

- (ii) a query store for storing database queries;
- (iii) a search tool for searching the query store against a constructed query; and
- (iv) query submission means for selecting between a constructed query and a query located against it by the search tool, and for submitting the selected query to a database.

Corporate databases are continually increasing in size and complexity. Over a period of time many queries will be submitted to these databases often repeating what was asked previously. Each query that is created incorporates implicitly some knowledge which could be useful for future similar queries. It has been recognised in making the present invention that it is possible to make this knowledge widely available across a community of users, for instance across a company.

Embodiments of the present invention make it possible not only to exploit a corporate database query knowledge but also to control queries submitted to large data repositories. That is, the combination of the query store and the means for constructing database queries allows structure to be imposed on new queries as well as allowing reuse of previous queries.

Advantageously, the query store and search tool use case based reasoning (CBR) principles and the means for constructing database queries does so to create a query as a case. This can be used to impose structure to be imposed on new queries.

CBR is a known technique which uses previous cases of problems and their solutions to solve new problems. A relevant publication is "Case-Based Reasoning: Experiences, Lessons and Future Directions" edited by D Leake and published in 1996 by AAAI Press, Menlo Park.

Embodiments of the present invention are particularly intended for use with structured query languages such as SQL. Intelligent query management according to embodiments of the present invention can then support a wide range of users with varying skill levels, from the complete SQL novice to the SQL expert, and can do it  
5 efficiently and effectively. Using a case-based approach, consistent and best practices can be encouraged across the organisation. Not only can the system itself improve over time as new cases are added but also the users' skill levels will increase due to inherent 'training' characteristics of the CBR systems. Furthermore, because experience is being captured the organisation will tend to be less vulnerable to the  
10 skill-loss effects of staff turnover.

From a technical viewpoint, CBR has also been found particularly appropriate. SQL has a well-defined and reasonably limited basic structure. This structure can be easily mapped into a case structure. For a target user base such as marketing, the requirements are often in the form of free-text. The natural language  
15 and text matching capabilities commonly found in CBR packages are therefore extremely useful.

In addition to storing a query, the results of the query can be stored. Preferably, therefore, embodiments of the present invention further provide a query results store and means for searching it. If a newly constructed query is identical or  
20 similar to a query already stored in the query store, then it may be that existing results can be used and neither the new query nor the stored query need be submitted. This saves precious central processing time in the database (Of course, if the stored results are too old then it may still be necessary for one of the queries to be selected and submitted in order to get fresh results). Preferably, the query store is provided in a  
25 cache.

A query construction tool will now be described as an embodiment of the present invention, by way of example only, with reference to the following figures, in which:

Figure 1 shows schematically the context in which the query construction tool sits in use;

Figure 2 shows the components of the query construction tool of Figure 1;

Figure 3 shows a flow chart of the overall process provided by the query construction tool; and

Figures 4, 5 and 6 show examples of user interface states presented to a user in use of the query construction tool.

Referring to Figure 1, the query construction tool 100 can be installed on processing capacity on any platform capable of supporting it. In this embodiment, the tool 100 is installed on a server 105 connected to a local area network (LAN) 110. Users have access from client workstations 115 which may be personal computers or may for instance primarily provide processing and interface capacity for running software processes normally installed elsewhere.

In order to support the tool 100, it is necessary to install an environment on the server 105 which supports case-based reasoning. An environment which was found particularly suited was ART\*Enterprise provided by Brightware Inc. Information on this environment can be obtained from the Brightware Inc. Website on the Internet located at "[http://www.Brightware.com/www\\_home.html](http://www.Brightware.com/www_home.html)". The environment was chosen because of its flexible CBR facilities integrated seamlessly with rule based and object oriented capabilities. Another important feature was its ability to create case-bases from existing data sources in various formats including Oracle, Dbase4, Excel and

Access. The environment also provides user interface construction facilities which allow demonstrations to be built quickly.

Other CBR environments could however be used in place of ART\*Enterprise, such as Inference. Inference however is not particularly well-suited. Inference is  
5 structured more for a "Help Desk " scenario and would require a lot of questions to be answered automatically, within the tool itself, in order to be useful. Also, Inference does not support direct access to existing Oracle data sources.

The server 105 is connected via the LAN 110 to the Internet 125 and can therefore be accessed by users over the Internet. Otherwise, users can use a LAN  
10 connection or a local connection to the server 105. The user's workstation 115 needs to be able to support an ART\*Enterprise client and the processing capability of at least a 486 processor is probably required for reasonable performance.

A database 120 which the tool 100 is used to submit queries to may be remote to the server 105, for instance accessible via the LAN 110, or could be local to the  
15 server 105, even for instance actually installed on the server 105.

The tool 100 comprises data as well as processing logic, the data being stored in a query store 130 and a results store 135. These stores may be external to the tool 100 or built into it. Such choices are well known to the person skilled in the art and are not further discussed herein.

20 Referring to Figure 2, in the same way it is not necessary that the database 120 to be queried should be sited as shown in Figure 1. As in distributed environments generally, it does not necessarily matter to a great extent where data is stored, as long as there is capacity and access and it may be preferable to site the database 120 to be queried on the server 105, together with the query construction  
25 tool 100.



The primary functional blocks of the query construction tool 100 are a query generator 200, results presentation 205 and management information processing 210. In support of the management information processing 210, there is also provided a behaviour database 230 which maintains count information with respect to queries in  
5 the query store 130.

#### **Overview Of CBR-Based Querying Process.**

In general, a user can use the query generator 200 to construct a new query and submit it to the query casebase 130 as a "case" in CBR terms. If there are similar queries already in the casebase 130, they will be returned to the user workstation 115  
10 where the user can select either to use their constructed query or an existing query. The selected query is then submitted to the main database 120. Results are returned to the user workstation 115 via the results presentation capability 205.

Alternatively, where there was an existing query, there may already be results stored in the results store 135. The user may choose not to submit either their  
15 constructed query or the existing query but just to view the stored results.

The user can start from nothing or use an existing query as the basis for describing the requirements for a new query. Once the user is happy with the description of a query they have constructed, then a case-based search of the casebase 130 is used to find existing queries which may meet their requirements.  
20 Any of the returned queries can be selected and browsed. If the results for a particular query exist they will have been stored in the results database 135 and the user may view them immediately. Alternatively the user may re-submit that query with or without modification. If a new query is sufficiently different from those stored in the past then it is added automatically to the casebase 130, whether it was constructed  
25 entirely from scratch or is an amended version of an existing query.

The count information maintained by the management information processing capability 210 in the behaviour database 230 can later be used to review usage patterns with respect to the query database 130 and the stored results database 135. In particular, there are three counts, these being "select" which increments when a query in the case base 130 is selected, "view" which is incremented when results stored in the stored results database 135 are accessed and "refresh" which is incremented when stored results in the stored results database 135 are updated.

Referring to Figure 3, the querying process may be as follows:

**STEP 300:** using the query generator 200, the user enters text description and/or any known SQL code to relevant fields of a query generator window to construct a query to present to the system. The entries made by the user constitute a case which can be searched against for matches in the casebase database 130. The tool can also construct a standard format SQL query to submit to the main database 120 from the entries made by the user simply by building the SQL query from the appropriate entries.

**STEP 305:** a case-based search is done against the presented query in the query casebase 130 and the most closely matching queries stored in the casebase 130 are returned.

**STEP 310:** queries stored in the casebase 130 carry an indicator in a "RESULTS-EXIST" field where the query has previously been run and there are results in the results database 135. At **STEP 310**, a check is made for each closely matching query which has been returned from the casebase 130 at **STEP 305** as to whether such results are already stored.

**STEP 315:** it may be that no closely matching queries were returned at **STEP 305**. In this case the user may want to start again, particularly if they are inexperienced and

need to run an existing query or need to have an existing query to modify so as to create a new query. The user is therefore offered the chance to start again at **STEP 300**, in which case the inexperienced user will probably try again to construct a new query that will bring an existing query at **STEP 305**. However, the experienced user  
5 may opt to proceed with the query they created at **STEP 300** or to base the query on a less closely matching query or even a dissimilar query found at **STEP 305**. This user will then be transferred to **STEP 325** below and offered the opportunity to edit their query before submission to the main database 120.

**STEP 320:** if results were located at **STEP 310**, being already stored for a closely  
10 matching query, the user is offered the opportunity to have them downloaded.

**STEP 330:** if the user selects to view the existing results, the closely matching query is submitted to the stored results database 135 and the results are downloaded to the user.

**STEP 325:** if the user selects not to view existing results at **STEP 320** for the closely  
15 matching query, the user is offered an opportunity to edit the closely matching query before moving to **STEP 330**, this time to submit the query to the main database 120. (If the user has not edited the closely matching query, this will simply have the effect of updating the search results for that query.)

**STEP 335:** any query which has been submitted to the main database 120 is  
20 reviewed in case it constitutes a new query. This is taken to be the case whenever the query submitted has less than a certain similarity threshold with any stored query in the casebase 130. This may be the situation for instance where the user has edited a closely matching query before submission to the main database 120.

**STEP 345:** new queries are automatically loaded to the casebase 130.

**STEP 340:** results from STEP 330 are returned to the user.

It can be seen from STEP 330 above that submission of a query to the stored results database 135 and the main database 120 can be treated similarly. In each case, the query being submitted is a SQL query. It is convenient that the stored  
5 results database 135 is located with the main database, simply being partitioned therefrom. That is, the stored results database 135 is built in table format and embedded in the main database 120.

Referring to Figures 2 and 3, it has been found convenient that, once the user has defined a query for submission to the main database 120, a results window 205 is  
10 presented to the user and the query is submitted directly to the main database 120 from the results window 205.

### **Other Capabilities**

In addition to the basic capabilities outlined above the following are provided:

- 15 • The user can choose to view a list of the five most frequently used queries. The user can select one of these to run directly or to use as the basis of a new query,
- The user can be presented with meta-data associated with an existing query. Meta-data is a known term for information associated with data, other than the data itself. The meta-data in this case includes for instance the original author,  
20 the date, query execution time, whether the results exist and if so, an identifier and size,
- The user has the option to run a subsequent query on the results returned or previously stored,
- The user has the option of updating the stored results database 135 with the  
25 results of rerunning an existing query.

### Case Structure, Matching and Retrieval

A major part of embodiments of the present invention is the definition of a case based on a SQL query. To define each case for the case base 130, a basic SQL statement is broken down by its structure, and parts of that structure are then used as  
5 fields for a case in the case base 130.

CBR matches cases against each other and different weights can be given to matches between different fields in a case. For instance, a field might be the "from clause" of a SQL query which determines which databases, or sections of databases, are to be searched. This might be given particularly high weighting so that a SQL  
10 query will only be selected as a similar query and offered to the user if it looks at the same tables of a database.

The structure of the cases is described in Table 1 below.

Feature	Match Type	Match Weight	Mis-match Weight
NOTES	Word	20	0
SELECT-CLAUSE	Word	25	0
FROM-CLAUSE	Character	45	0
WHERE-CLAUSE	Character	10	0
GROUP-BY-CLAUSE	Word	5	0
HAVING-CLAUSE	Word	1	0
ORDER-BY-CLAUSE	Word	1	0
ROWS_SELECTED	#	#	#
QUERY_TIME	#	#	#
CREATOR	#	#	#
RESULTS_EXIST	#	#	#
RESULTS_REF	#	#	#
DATE_CREATED	#	#	#

15 Table 1. Case Structure

The NOTES feature is free text describing the data that should be returned from the query to be written. The main features are then those that correspond to the structure of a simple SQL query. The items in italics are the meta-data. This is additional information which may have an impact on the user's final choice but is not part of the casebase index. In addition each case has a unique identifier (not shown in the table).

The choice of match types and weights is optional. In the example of Table 1, they are based on knowledge about the domain of SQL querying and initial experimentation. For example, the most important feature is the FROM-CLAUSE because this determines the specific tables from which the data will be retrieved. It will be of "character" type because of the manner of identifying tables in the database to be searched. In this particular example, queries are not penalised for not matching against a feature value and therefore all mis-match weights are zero.

An example case is as follows:

Feature	Value
NOTES	Top ten business for ISDN2
SELECT-CLAUSE	distinct (a.local_number) , tally, direction, position, s.site_number, business_name
FROM-CLAUSE	CCBA_ISDN_2 a, sites s
WHERE-CLAUSE	a.site_number = s.site_number and position <= 10
GROUP-BY-CLAUSE	(empty)

HAVING-CLAUSE	(empty)
ORDER-BY	(empty)
ROWS-SELECTED	10
QUERY-TIME	00:00:03
CREATOR	RTS
RESULTS-EXIST	YES
RESULTS-REF	TOPTENPROD
DATE-CREATED	2-Feb 98 16:00

Table 2. Example Case

Table 2 shows the fields of a SQL query to locate the top ten communications sites generating traffic on ISDN2 lines (Integrated Services Digital Network lines based on the capacity of two conventional speech telephony lines).

- 5        The matching process between queries is a straight forward use of CBR and applies the in-built capabilities of the environment as described in ART\*Enterprise Brightware Inc, 1996, *ARTScript Programming Guide 3, Rules & CBR*. Each feature is compared and scored according to its type: *string*; *word*; and *character*. These are combined taking into account the weights to give an overall feature score, or similarity
- 10    score, for the case. A threshold of 0.4 is set above which a case must score to be retrieved. (The feature score can also be used to determine automatically when a query is sufficiently different from stored queries to be added to the casebase database 130. In these circumstances, a score of less than 0.4 means that the query will be added to the casebase database 130.) The maximum number of cases
- 15    retrieved is set to 10. These values are arbitrary but have been found to be convenient. Other values may be preferred for instance where an embodiment of the invention is applied to quite a different domain or for a different purpose.

The feature score for numeral matching is generated in a known manner, as follows:

$$feature - score_{f,i} = mw_{f,i} - \frac{|cv_{f,i} - pcv_{f,i}|}{mdev_{f,i}} (mw_{f,i} - mmw_{f,i})$$

where:

- $mw_{f,i}$  is the match weight of feature  $f$  for case  $i$ ;
- $mmw_{f,i}$  is the mismatch weight of feature  $f$  for case  $i$ ;
- 5  $mdev_{f,i}$  is the match deviation of feature  $f$  of case  $i$ ;
- $cv_{f,i}$  is the numeric value for feature  $f$  of case  $i$ ; and
- $pcv_{f,i}$  is the numeric value for feature  $f$  of the presented case.

The feature score for text matching is generated in known manner. The formula is as  
10 follows:

$$feature - score_{f,i} = mmw_{f,i} + \frac{msf_{f,i}}{tsf_f} (mw_{f,i} - mmw_{f,i})$$

where:

- $mmw_{f,i}$  is the match weight of feature  $f$  for case  $i$ ;
- 15  $mw_{f,i}$  is the mismatch weight of feature  $f$  for case  $i$ ; and
- $msf_{f,i}$  is the number of matching sub-features of feature  $f$  for case  $i$ .

Domain specific aliases (i.e. different words which have the same or similar meaning) may be taken into account during this calculation. The result is that the quality of the  
20 text matching is improved because the algorithm matches on meaning (semantics) rather than raw text (syntax). Typically, this is achieved by replacing each word by a symbol as determined by reference to a thesaurus that maps different words having the same meaning for a given domain to the same symbol. Data storage can be allocated to provide means for storing the thesaurus. The thesaurus functionality may  
25 be provided by a knowledge base (not shown) which includes domain specific knowledge and means for mapping that knowledge to non-domain specific knowledge.

A number of data areas are defined for embodiments of the present invention:

- 30 • Database 120

The location of the data on which the queries will be run



- Query Casebase 130

The casebase containing the past queries

- Stored Results 135

The data sets returned by the past queries

- 5 • Behaviour Database 230

The functions of most of these have been discussed above. The behaviour database 230 however stores the values of counters which record the usage of each case. The counters are; (i) 'Select', which increments when a query is selected, (ii) 'View', which  
10 increments when the stored results are accessed and, (iii) 'Refresh', which increments if stored results are updated.

Each count can be provided as a field associated with a result in the results database 135 or with a query in the case base 130. Each time a count is incremented, it is incremented by updating the count field. Management information  
15 can then be presented by searching the results database 135 and/or the query case base 130 to locate results or queries with count fields putting them above a selected threshold, such as in the top five or top ten of the whole database 135, 130 in each case. However, as shown, the counts are maintained in an additional data store, the behaviour database 230, where queries or results can be associated with their  
20 relevant count by their position in an array or by being allocated respective identifiers against which the counts are maintained.

One way of designing the behaviour database 230 is to structure the count data as cases. It can then be searched by running a dummy case in which all the counters are set to maximum and searching for a selected number of most closely  
25 matching cases, such as the "Top 5" cases. That is, a case match is done just on the count fields.

Figures 4 to 6 show the three main windows presented to the user in various statuses of the casebased query construction tool 100.

Referring to Figure 4, the Top 5 Queries window 410 shows the most 'popular' queries based on their usage as defined by the values of the counters of the behaviour database 230. This window is one of several windows that might be presented to the user for accessing the behaviour database 230. The user can select a query listed on the Top 5 Queries window 210 to submit directly to the main database 120, to use as a presentation case (optionally modified) for a CBR search, or to modify and submit i.e. as input to STEP 300.

Referring to Figure 5, the Query Generator window 400 is the main search and construction window. The user enters descriptive values into the fields (not shown) associated with each feature. At any time the user can present the query to the casebase 130 or send the query for submission to the main database 120. Buttons present on the window are;

- TOP 5: the function attached to this button causes the Top 5 window to appear
- SEARCH: this function invokes a casebase match and retrieval. The results of the search appear in the drop-down-list-box at the top of the window.
- CLEAR: clears the entries in the fields
- ALTER CASE: allows the user to modify the contents of a case
- CANCEL: closes the window
- SUBMIT: executes the SQL in the window, after checking the validity of the code. If the SQL submitted is sufficiently dissimilar from any contained within the casebase 130 then a new case is created automatically without the user knowing. The returned results are presented in the Results window.

Referring to Figure 6, the Results window 405 displays the final SQL 600 as generated from the Query Generator window and the returned results 605. The SQL can be modified and resubmitted directly from this window 405 allowing a follow-up query on previously returned results.

5 (A window which is not shown here is the database login window. This is of standard type however.)

In embodiments of the present invention in use, the user may enter values for as many case features as required. If just the notes feature is specified then queries can be found without any knowledge of SQL or the database schema. This allows  
10 complete SQL novices to retrieve queries. Of course, it may be that unless an exact match is found then such a complete novice will not be able to proceed because modifications to the query would be required. Whether this is a problem depends upon the circumstances. It could be imagined that the novices would deal with straightforward queries and pass those that needed changes to more senior users.

15 Improvements which the person skilled in the art might wish to make for a particular use of embodiments of the present invention are as follows:

- Tuning of the choice of feature match types and weights;
- Inclusion of a full domain specific thesaurus to cope with aliases within and across databases;
- 20 • Development of use of the meta data, for example, prediction of query times or deriving a 'cost' metric for the query. Such development is discussed in a paper by Liu, L., Pu, C. 1997, entitled *A Metadata Based Approach to Improving Query Responsiveness*, (<http://computer.org/conferen/proceed/meta97/papers/liu/lingliu-full.html>);
- 25 • Techniques for supporting time/resource constrained queries;

- Long term management of the casebase, including policies for the creation, deletion and indexing of cases;
  - Making use of other types of meta-data such as information in the schema of the database being queried;
- 5 • Making more use of the behaviour data, for example, to monitor user skill/familiarity levels for intelligent allocation of query requests, or to detect trends across users at different times of the year to plan resources optimally; and
- Acceptable ways of obtaining user feedback on how useful the query was. It is known that users become frustrated when continually forced to enter supplementary
- 10 information not directly relevant to achieving their task at hand.

## CLAIMS

1. A database access tool comprising:
  - i) means for constructing database queries;
  - 5 ii) a query store for storing database queries;
  - iii) a search tool for searching the query store against a constructed query; and
  - iv) query submission means for selecting between a constructed query and a query located against it by the search tool, and for submitting the selected query to a database.
- 10 2. A tool according to Claim 1 wherein the means for constructing database queries comprises user input means for loading data to at least one data field in a database query and the search tool comprises means to calculate a similarity factor between data fields in database queries stored in the query  
15 store and at least one data field in a newly constructed database query.
3. A tool according to Claim 2 wherein the query store and the search tool are constructed according to use of case based reasoning (CBR) and the means for constructing database queries does so to construct a query as a case.
- 20 4. A tool according to Claims 2 or 3 wherein management information data is collected in use of the tool to submit queries to the database, the tool further comprising means for collecting said management information, structuring the management information for a respective query in the same manner as a

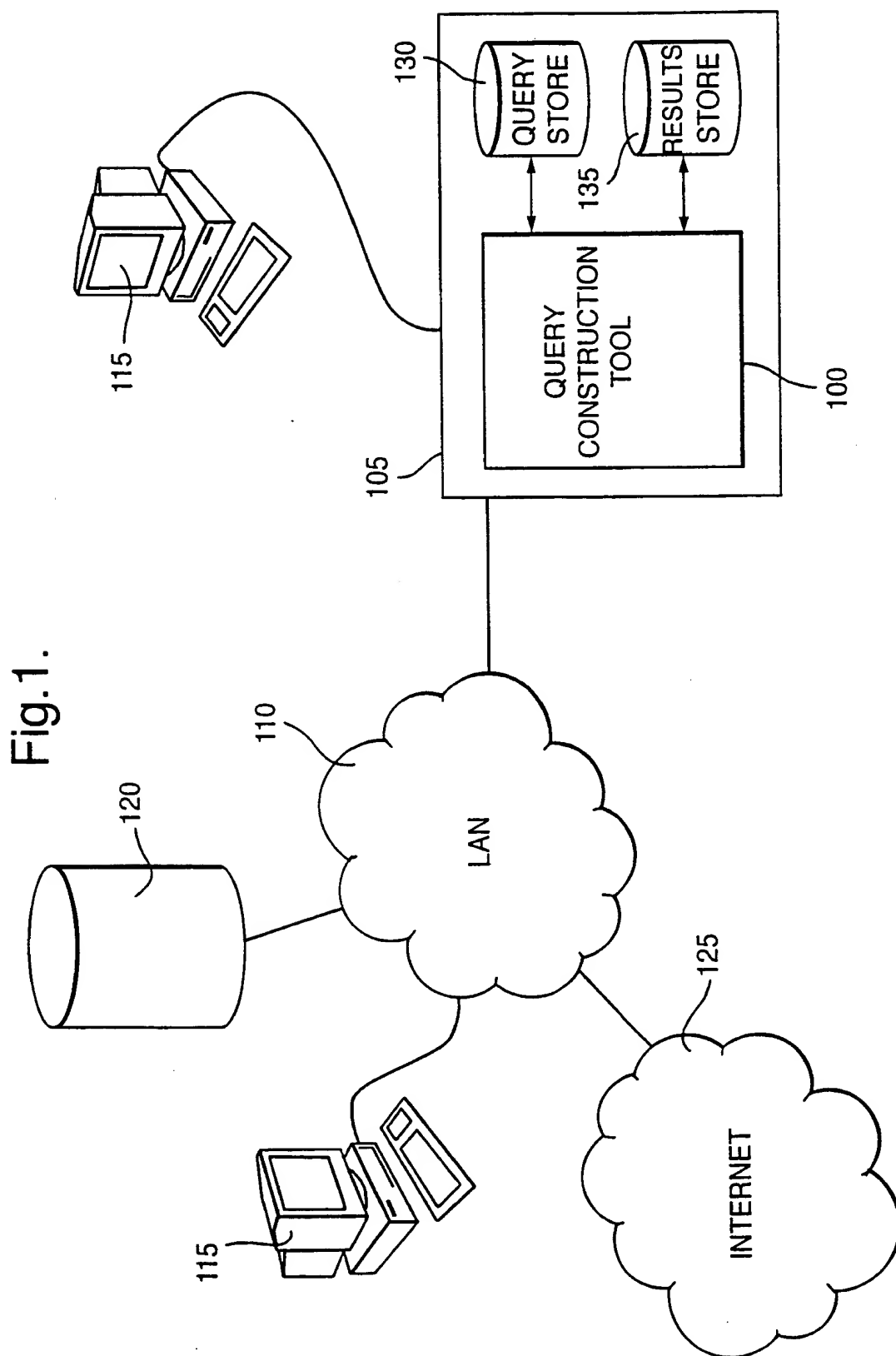
database query constructed by the tool and loading the structured management information to a management information data store, said management information data store being searchable by means of the search tool.

5

5. A database access tool according to any preceding claim, further comprising a data store for storing results associated with previous database queries and wherein the search tool is further operable, when a located query is identical or similar to the constructed query, to return the results stored in the data store that are associated with the located query.

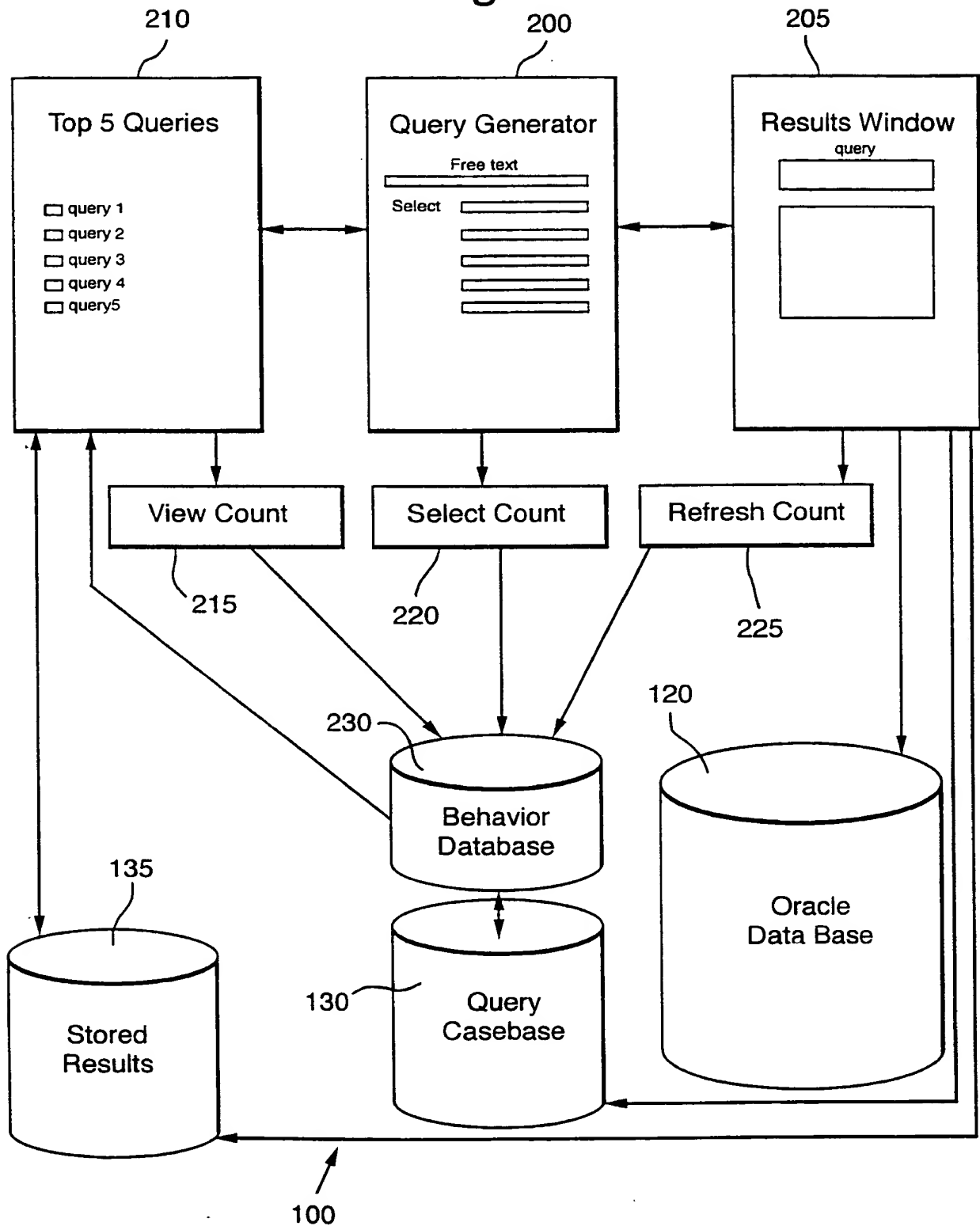
10

1/5



2/5

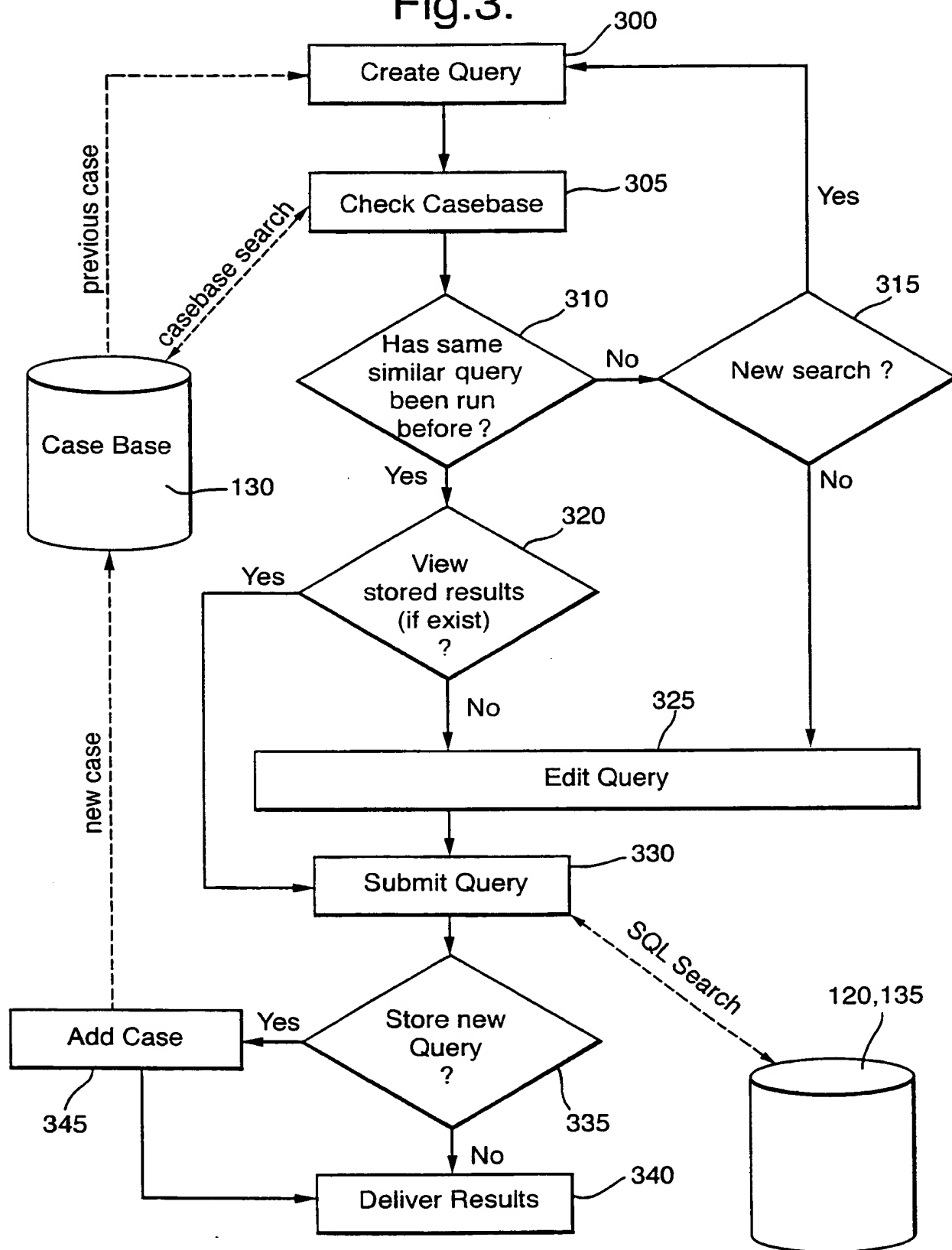
Fig.2.





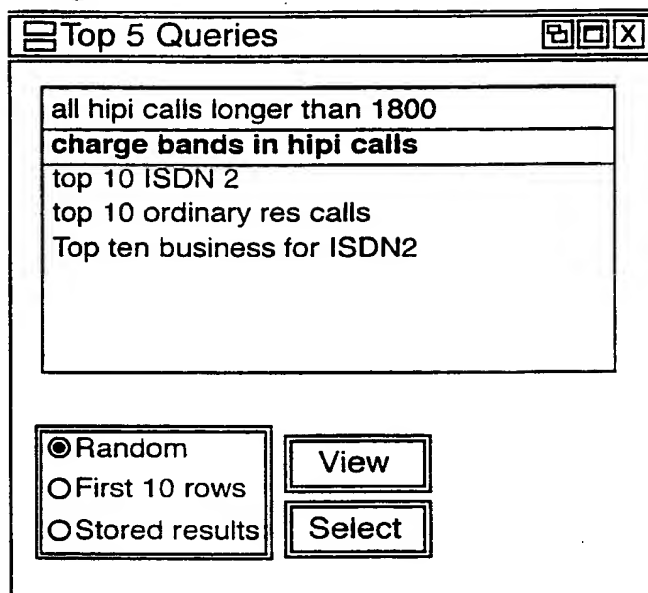
3/5

Fig.3.



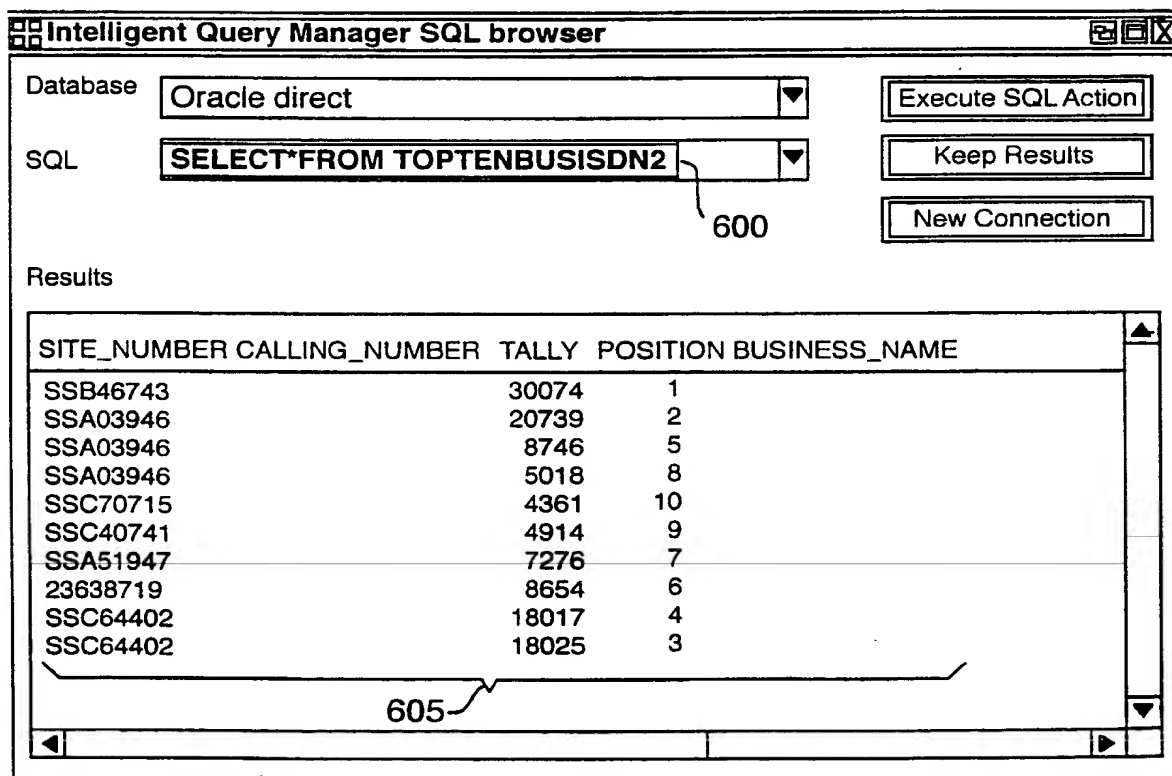
4/5

Fig.4.



410

Fig.6.



405

5/5

Fig.5.

Multi User Knowledge Capture Administrator

90: Top ten business for ISDN2 ▼

Free Text Description

Top ten business for ISDN2

Top 5

Search

Clear

Alter Case

Cancel

Meta Data

How Selected Query Time 00:00:00 Results Exist YES

Creator REG **Results Ref** TOPTENPROD Number 4

Date:-----

400

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 99/01278

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 6 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	"SQL STATEMENT LIBRARY LIST" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 36, no. 1, 1 January 1993 (1993-01-01), pages 82-83, XP000333783 ISSN: 0018-8689 the whole document ---	1
A	"STORING QUERY OBJECTS BY TYPE IN A RELATIONAL DATABASE" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 32, no. 4B, 1 September 1989 (1989-09-01), page 14 XP000067092 ISSN: 0018-8689 the whole document --- -/--	1

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

2 August 1999

Date of mailing of the international search report

09/08/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2

NL - 2280 HV Rijswijk

Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,

Fax: (+31-70) 340-3016

Authorized officer

Deane, E

# INTERNATIONAL SEARCH REPORT

Int ernational Application No

PCT/GB 99/01278

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>"STORE SEARCH RESULTS FOR LATER REFERENCE"</p> <p>IBM TECHNICAL DISCLOSURE BULLETIN, vol. 32, no. 10B, 1 March 1990 (1990-03-01), page 210 XP000097865</p> <p>ISSN: 0018-8689</p> <p>the whole document</p> <p>---</p>	5
A	<p>PATENT ABSTRACTS OF JAPAN</p> <p>vol. 098, no. 005, 30 April 1998 (1998-04-30)</p> <p>&amp; JP 10 011339 A (HITACHI LTD), 16 January 1998 (1998-01-16)</p> <p>abstract</p> <p>&amp; US 5 905 983 A</p> <p>abstract</p> <p>-----</p>	5

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/GB 99/01278

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
JP 10011339 A	16-01-1998	US 5905983 A	18-05-1999